The Analysis of Interlaboratory Study Data

Mark G. Vangel
Statistical Engineering Division
National Institute of
Standards and Technology
100 Bureau Drive, Stop 8980
Gaithersburg, MD 20899-8980

May 7, 1999

Outline

- Introduce interlaboratory studies, with motivation and examples.

- Review work of W.J. Youden and John Mandel, which has had a substantial influence on practice.
  - Youden plots
  - Row-Linear Models

- Present some recent research, focusing mostly on the heteroscedastic one-way random model.
  - Mandel-Paule procedure as approximate REML.
  - ML and profile likelihood analysis.
  - Bayesian inference.

Interlaboratory Studies: The Scenario

- Each of p laboratories makes repeated measurements of m quantities (perhaps corresponding to different concentrations of a chemical analyte).

- The number of measurements made can differ among the laboratories.

- The measurement variability may depend on the material being measured (perhaps as an increasing function of concentration or level).

- The within-laboratory variabilities may differ (often, though, they are assumed to be equal).

Interlaboratory Studies:
Some questions

- How should one estimate 'consensus' values of the quantities measured?

- What is the between-laboratory variability (*reproducibility*)?

- What is the within-laboratory variability (*repeatability*)? How do they compare?

- How should we look for outliers?

Why Interlaboratory Studies?

- Interlaboratory studies are primarily performed for one of two reasons:

  1. Validating a measurement method or standard material

  2. Assessing the proficiency of measurement laboratories.

Examples of Both Types of Studies

- An enzymatic-gravimetric method is developed for measuring the dietary fiber in foods. Standardized samples of foods are prepared, and distributed to various testing laboratories, who measure the concentrations using the proposed method.

- The National Research Council of Canada and NOAA together conduct interlaboratory comparisons to evaluate the proficiency of test laboratories at determining concentrations of trace elements in marine biological tissues. Homogeneous materials (e.g. oyster tissue, marine sediments) are distributed among various laboratories, who return data on several trace elements (e.g., arsenic).

Evaluating an Analytical Method for Dietary Fiber
Li and Cardozo (1994)
J. Of AOAC Int., 77, p. 689

Nine labs each measures fiber in six foods, in *blind duplicates.*

Sample	Laboratory			
	1	2	$\cdots$	9
Apples	12.44	12.87	$\cdots$	12.08
	12.48	13.20	$\cdots$	12.38
Apricots	25.05	27.16	$\cdots$	25.31
	25.58	26.29	$\cdots$	25.43
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
FIBRIM	74.07	76.55	$\cdots$	73.96
	75.01	78.36	$\cdots$	74.24

Unequal Variances for Fiber Data

Arsenic in Oyster Tissue (NIST Standard Reference Material 1566a)

Interlaboratory Study
Methodology at NBS
W.J. Youden (1948-1972)

- **W.J. Youden** A chemist with statistical training, Youden was very active in the Association of Official Analytical Chemists. He wrote a manual of procedures to be used by members of this association, in which he introduced the influential concept of the **Youden Plot**.

Interlaboratory Study
Methodology at NBS
John Mandel (1948-   )

- **John Mandel** Also a chemist with statistical training, Mandel's career has been devoted to work in interlaboratory studies, and to understanding measurement as a process. His most influential contributions center on the **Row-Linear Model** for two-way tables in interlaboratory studies, and on a novel approach to single material studies. He has long been an influential member of the American Society for Testing and Materials.

The Youden Plot

- Prepare a material in pairs of *blocks*. Each block is as homogeneous as possible, and is divided into samples, with one sample sent to each laboratory.

- The blocks need not be identical, but should be close in the mean level of the quantity being measured, so that the measurement variances can be assumed to be equal.

- Several pairs of blocks can be used to cover a wide range of concentrations.

Youden Plot: Linear Model

- Let the data from a pair of blocks be x_{ij}, $i = 1, 2$ and $j = 1, \ldots, p$.

- We assume that

$$x_{ij} = \mu_i + b_j + e_{ij}$$

  ($i = 1, 2$: Samples; $j = 1, \ldots, p$: Labs)

- Laboratory effect: b_j, different for each lab, normal with mean 0 and variance σ_b^2.

- Measurement errors: e_{ij}, assumed to identically distributed normal with mean 0 and common variance σ_e^2.

- Plot the x_{1j}s against the x_{2j}s. A circular pattern indicates no laboratory effect; an ellipse indicates a laboratory effect.

- Outlying labs can be qualitatively determined.

Statistical Inference

- Points (x_{1j}, x_{2j}) equicorrelated normal with mean (μ_1, μ_2).

- Contours of bivariate distribution are elliptical, with semimajor axis

$$a \propto \sqrt{2\sigma_b^2 + \sigma_e^2}$$

and semiminor axis

$$b \propto \sigma_e$$

- Variance component estimates follow from the independent sample variances

$$\text{Var}\left(\frac{x_{1j} + x_{2j}}{\sqrt{2}}\right) \sim (2\sigma_b^2 + \sigma_e^2)\chi_{p-1}^2$$

$$\text{Var}\left(\frac{x_{1j} - x_{2j}}{\sqrt{2}}\right) \sim \sigma_e^2 \chi_{p-1}^2$$

- The ratio of these estimates is the F-ratio for testing $\sigma_b^2 = 0$.

- Confidence and prediction regions are strightforward to calculate.

Youden Plot for Proficiency Study: Arsenic in Marine Tissue

Two-Way Tables

- The typical data structure for an interlaboratory study is a two-way table, although sometimes (as above) the data are analyzed one material at a time.

- One way to model such data is a two-way ANOVA with interaction:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk}$$

where

  - α_i, $i = 1, \ldots, p$ is the *laboratory effect* (perhaps random)
  - β_j, $j = 1, \ldots, m$ is the *material effect* (fixed)
  - γ_{ij} is the lab/material interaction
  - e_{ijk}, $k = 1, \ldots, n_{ij}$ is the measurement error, with variance which probably depends on material.

Mandel's Approach to Two-Way Tables

- Typically, one sees unequal error variances for different materials, and often nonadditivity as well.

- Transforming the data can help, but Mandel argues that this is not appropriate since there are multiple variances in the model.

- Mandel's approach consists of

  1. Estimating the within variance separately for each material, and then reducing the data to cell means.

  2. Estimating the row effects, column effects, and interaction.

  3. Regressing the estimated interaction against the column (material) effects. This results in a decomposition of the interaction into a part due to slopes among labs, and a residual.

Calculations for the Row-Linear Model

- Error variances:

$$s_j^2 = \frac{\sum_{i=1}^{p} \sum_{k=1}^{n_{ij}} (y_{ijk} - \bar{y}_{.j.})^2}{\sum_{i=1}^{p} (n_{ij} - 1)}$$

- Effects:

$$
\begin{aligned}
\widehat{\mu} &= \bar{y}_{...} \\
\widehat{\alpha}_i &= \bar{y}_{i..} - \bar{y}_{...} \\
\widehat{\beta}_j &= \bar{y}_{.j.} - \bar{y}_{...} \\
\widehat{\gamma}_{ij} &= \bar{y}_{ij.} - \widehat{\mu} - \widehat{\alpha}_i - \widehat{\beta}_j
\end{aligned}
$$

$$(1)$$

- Row-linear model for interaction:

$$\widehat{\gamma}_{ij} = b_i \widehat{\beta}_j + h_{ij}$$

where b_i is the least-squares slope for the ith lab., and h_{ij} is the part of the interaction not explained by the linear regression.

Laboratory Linear Regressions
for the Fiber Data

- If we do these linear regressions for the fiber data, we find some very significant slopes. But also some insignificant ones.

- The significant slopes are strongly influenced by the Fibrim data.

Lab.	$\widehat{b}_i$	$s_{\widehat{b}_i}$	$\widehat{b}_i / s_{\widehat{b}_i}$	P-Value
1	-0.0256	0.0075	-3.4129	0.0270
2	0.0140	0.0044	3.1827	0.0334
3	0.0123	0.0207	0.5940	0.5845
4	-0.0095	0.0115	-0.8274	0.4545
5	-0.0005	0.0058	-0.0814	0.9391
6	-0.0504	0.0137	-3.6760	0.0213
7	0.0686	0.0064	10.7495	0.0004
8	0.0183	0.0120	1.5239	0.2022
9	-0.0272	0.0041	-6.5961	0.0027

The Row-Linear ANOVA Table

We can write Mandel's model as:

$$\bar{y}_{ij.} = \bar{y}_{i..} + (b_i + 1)(\bar{y}_{.j.} - \bar{y}_{...}) + h_{ij}$$

Some refer to this as Mandel's 'bundle-of-lines'.

The ANOVA table is

Rows	$p - 1$	$m \sum_i (\bar{y}_{i..} - \bar{y}_{...})^2$
Columns	$m - 1$	$p \sum_j (\bar{y}_{.j.} - \bar{y}_{...})^2$
Interaction	$(p-1)(m-1)$	$\sum_{ij} (y_{ij} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y})^2$
Slopes	$p - 1$	$\sum_{ij} b_i^2 (\bar{y}_{.j.} - \bar{y}_{...})^2$
Remainder	$(p-1)(m-2)$	$\sum_{ij} h_{ij}^2$

Row-Linear ANOVA Table for Dietary Fiber Data

Source	SS	df	F-Ratio
Labs.	64.48	8	23.08
Foods	12447871.91	5	7127822.96
Interaction	36.04	40	2.58
Slopes	24.86	8	8.90
Resid.	11.18	32	

P-Value for Slopes: 2.5×10^{-6}

Single-Material Interlaboratory Model: One-Way, Unbalanced, Heteroscedastic Random-Effects ANOVA

- Laboratory sample means $\bar{x}_{i\cdot}$ distributed independently normal with mean μ and variance $\sigma^2 + \tau_i^2$, where $\tau_i^2 = \sigma_i^2/n_i$.

- Expected mean for ith laboratory is also normal, with mean μ and variance σ^2.

- Sufficient statistics $\bar{x}_{i\cdot}$ and $t_i^2 = s_i^2/n_i$.

If x_{ij} denotes the jth measurement from the ith lab, then

$$x_{ij} = \mu + b_i + e_{ij},$$

where $b_i \sim \mathsf{N}(0, \sigma^2)$ and $e_{ij} = \mathsf{N}(0, \sigma_i^2)$; mutually independent.

One-Way Models in Interlaboratory Studies: The Mandel-Paule Estimator J. of Research of the NBS (1982)

- For arbitrary positive weights $\{w_i\}_{i=1}^{k}$, weighted mean is

$$\tilde{\mu} = \frac{\sum_{i=1}^{p} w_i \bar{x}_{i.}}{\sum_{i=1}^{p} w_i}.$$

- *Mandel-Paule* estimate, μ_{MP}, of μ is the weighted mean $\tilde{\mu}$ for which

$$w_i \equiv \frac{1}{\tilde{\sigma}^2 + t_i^2}$$

where $\tilde{\sigma}^2$ is the root (if any) of

$$Q = \sum_{i=1}^{p} w_i (\bar{x}_{i.} - \tilde{\mu})^2 = p - 1$$

- Note: Q is convex decreasing on $[0, \infty)$, and $Q \sim \chi_{p-1}^2$ if

$$w_i = \omega_i \equiv \frac{1}{\sigma^2 + \tau_i^2}$$

The Mandel-Paule Algorithm and ML/REML

Maximum-Likelihood for a linear model

$$Y = X\beta + e,$$

where $e \sim \mathsf{N}(0, \Sigma)$ is equivalent to minimizing of $|\Sigma|$, subject to

$$(y - X\widehat{\beta})^T \Sigma^{-1} (y - X\widehat{\beta}) = n \qquad (1)$$

where $\widehat{\beta}$ is the GLS estimate of β, and n is the number of observations.

For our one-way model, if the σ_i^2 are replaced by s_i^2, then (1), an equation in σ^2 alone, is

$$\sum_{i=1}^{p} w_i (\bar{x}_{i.} - \tilde{\mu})^2 = p.$$

Had REML been used, rather than ML, then the p on the RHS above would be a $p - 1$, *precisely* Mandel and Paule's equation.

A Problem With a Long History: Cochran's Publications on Combining Experiments

- (1937), "Problems Arising in the Analysis of a Series of Similar Experiments".

- (1938), "The Analysis of Groups of Experiments", (with F. Yates).

- (1954), "The Combination of Estimates From Different Experiments".

- (1980), "Summarizing the Results of a Series of Experiments".

- (1981), "Estimators for the One-Way Random Effects Model With Unequal Error Variances", (et. al., posthumous).

Maximum Likelihood
(Cochran, 1937)

Let $\omega_i = 1/(\sigma^2 + \tau_i^2)$, $\nu_i = n_i - 1$, and determine $\hat{\sigma}$, $\hat{\tau}_i^2$, and $\hat{\mu}$ to satisfy

$$(A_i) \quad \omega_i - \omega_i^2(\bar{x}_{i.} - \mu)^2 + \nu_i\left(\frac{1}{\tau_i^2} - \frac{t_i^2}{\tau_i^4}\right) = 0$$

$$(B) \quad \boxed{\sum_{i=1}^k \omega_i^2(\bar{x}_{i.} - \mu)^2 = \sum_{i=1}^k \omega_i}$$

$$(C) \quad \mu = \frac{\sum_{i=1}^k \omega_i \bar{x}_{i.}}{\sum_{i=1}^k \omega_i}$$

Note that (B) may have multiple roots. Cochran (1937) proposed setting $\tau_i^2 = t_i^2$ and solving (B) for σ^2, then using (C).

The Loglikelihood Function: A Better Parametrization

Define weights by

$$\gamma_i \equiv \frac{\sigma^2}{\sigma^2 + \tau_i^2}$$

The loglikelihood becomes

$$
\begin{aligned}
2\ell \;=\;& \sum_{i=1}^{p} n_i \log\left(\frac{\gamma_i}{\sigma^2}\right) \\
& - \sum_{i=1}^{p} \frac{\gamma_i}{\sigma^2}\left[(x_i - \mu)^2 + \frac{\nu_i t_i^2}{1 - \gamma_i}\right] \\
& - \sum_{i=1}^{p} \nu_i \log(1 - \gamma_i) + K.
\end{aligned}
$$

Differentiate this with respect to parameters μ, σ^2 and $\gamma_i, i = 1, \ldots, p$.

ML Equations

$$\mu = \frac{\sum_{i=1}^{p} \gamma_i \bar{x}_{i.}}{\sum_i \gamma_i} = \frac{\sum_{i=1}^{p} \omega_i \bar{x}_{i.}}{\sum_i \omega_i}$$

$$\sigma^2 = \frac{\sum_{i=1}^{p} \gamma_i \left[(\bar{x}_{i.} - \mu)^2 + \frac{\nu_i t_i^2}{1 - \gamma_i} \right]}{\sum_{i=1}^{p} n_i}$$

$$\gamma_i^3 - (a_i + 2)\gamma_i^2 +$$
$$[(n_i + 1)a_i + (n_i - 1)b_i + 1]\,\gamma_i$$
$$-n_i a_i = 0$$

where

$$a_i \equiv \frac{\sigma^2}{(\bar{x}_{i.} - \mu)^2}$$

and

$$b_i \equiv \frac{t_i^2}{(\bar{x}_{i.} - \mu)^2}.$$

Result #1:
Monotone Convergence to Stationary Points of the Likelihood

- For any starting values μ_0, σ_0^2, maximize the likelihood over the weights by solving the cubics. (If there are multiple real roots, choose the one which causes the biggest increase in the likelihood.)

- Let

$$\sigma_1^2 = \frac{\sum_{i=1}^{p} \gamma_i \left[(\bar{x}_{i.} - \mu)^2 + \frac{\nu_i t_i^2}{1-\gamma_i} \right]}{\sum_{i=1}^{p} n_i}$$

$$\mu_1 = \frac{\sum_{i=1}^{p} \gamma_i \bar{x}_{i.}}{\sum_{i=1}^{p} \gamma_i}$$

  solve for new weights, and iterate.

- This iteration, *regardless of starting values*, always converges to a stationary point of the likelihood, and *increases the likelihood at each step.*

Result #2:
Location of Stationary Values of the Likelihood

- At a stationary point of the likelihood,

$$\widehat{\sigma}^2 = \frac{\sum_{i=1}^{p} \gamma_i^2 (\bar{x}_{i.} - \mu)^2}{\sum_{i=1}^{p} \gamma_i}$$

  hence

- *All* of the stationary points of the likelihood $\widehat{\mu}$ and $\widehat{\sigma}$ are within the rectangle in the (μ, σ) plane given by

$$\min_i(\bar{x}_{i.}) \leq \tilde{\mu} \leq \max_i(\bar{x}_{i.})$$

  and

$$0 \leq \tilde{\sigma} \leq \max_i(\bar{x}_{i.}) - \min_i(\bar{x}_{i.}).$$

- After the appropriate location-scale transformation of the data, it is only necessary to search the unit square in the (μ, σ) plane for stationary values.

Lab. 6 an Outlier for Apricot Data

Outlier Labs. for Cabbage Data

Four Local Maxima for Fibrim Likelihood

Result #3:
Location of the Roots of Cubic Equations for Weights (γ_i)

- Each cubic likelihood equation has one or three roots $\gamma_i \in [0, 1]$.

- A necessary condition for three roots is that

$$(\bar{x}_{i.} - \mu)^2 \geq \max(\sigma^2/q_i, t_i^2/h_i),$$

where

$$
\begin{aligned}
q_i &= -2 - 6\sqrt{n_i} \sin\left\{\frac{1}{3}\left[\sin^{-1}\left(\sqrt{\frac{n_i - 1}{n_i}}\right) - \frac{\pi}{2}\right]\right\} \\
&= \frac{8}{27n_i} + O(n_i^{-2})
\end{aligned}
$$

and

$$h_i = \frac{(1 - q_i)^3}{27(n_i - 1)} = \frac{1}{27n_i} + O(n_i^{-2}).$$

- These values q_i and h_i are the smallest for which this is necessary.

A Comment on Homoscedastic Models

- If we require that the 'within' variances be equal, than we still have the likelihood equations

$$\mu = \frac{\sum_{i=1}^{p} \gamma_i \bar{x}_{i.}}{\sum_i \gamma_i}$$

$$\sigma^2 = \frac{\sum_{i=1}^{p} \gamma_i \left[(x_i - \mu)^2 + \frac{\nu_i t_i^2}{1-\gamma_i} \right]}{\sum_{i=1}^{p} n_i}$$

- The weights can be parametrized as

$$\gamma_i = \frac{u}{u + (1 - u)\frac{n_1}{n_i}},$$

for $0 \leq u < 1$. Maximizing the likelihood reduces to maximizing with respect to u. In particular, all of the stationary points of the likelihood must be on a curve in the (μ, σ) plane, and one need not be concerned about negative solutions for the variances.

Hierarchical Model With Noninformative Priors

$i = 1, \ldots, p$ indexes laboratories

$j = 1, \ldots, n_i$ indexes measurements

$$
\begin{aligned}
p(x_{ij}|\delta_i, \sigma_i^2) &= \mathsf{N}\,(\delta_i, \sigma_i^2) \\
p(\sigma_i) &\propto 1/\sigma_i \\
p(\delta_i|\mu, \sigma^2) &= \mathsf{N}\,(\mu, \sigma^2) \\
p(\mu) &\propto 1 \\
p(\sigma) &\propto 1
\end{aligned}
$$

Posterior given $\sigma = 0$, $p \geq 1$

Given $\sigma = 0$, then the posterior distribution of the consensus mean μ is proportional to a product of scaled t-densities:

$$p(\mu|\{x_{ij}\}|\sigma = 0) \propto \prod_{i=1}^{p} \frac{1}{t_i} T'_{n_i-1}\left(\frac{\bar{x}_{i.} - \mu}{t_i}\right)$$

The General Case: $\sigma \geq 0$

In general, $p(\mu|\sigma, \{x_{ij}\})$ is proportional to a *product* of the distributions of the random variables

$$U_i = \bar{x}_{i.} + \frac{\sqrt{n_i}}{s_i} T_{n_i-1} + \sigma Z,$$

where T_{n_i-1} is a t-distributed random variable with $n_i - 1$ degrees of freedom, Z is distributed $\mathsf{N}(0, 1)$, and T_{n_i-1} and Z are independent.

A Useful Probability Density

Let T_ν and Z denote independent Student-t and standard normal random variables, and assume that $\psi \geq 0$ and $\nu > 0$. Then

$$U = T_\nu + Z\sqrt{\frac{\psi}{2}}$$

has density

$$f_\nu\left(u; \psi\right) \equiv \frac{1}{,_{\nu/2}\sqrt{\pi}} \int_0^\infty \frac{y^{(\nu+1)/2-1}e^{-y\left[1+\frac{u^2}{\psi y+\nu}\right]}}{\sqrt{\psi y + \nu}}dy.$$

Posterior of (μ, σ)

- Assume $\delta_i \sim N(\mu, \sigma^2)$, $\sigma \sim p(\sigma)$,
  $p(\mu) \propto 1$, $p(\sigma_i) \propto 1/\sigma_i$.

- Then the posterior of (μ, σ) is

$$p(\mu, \sigma | \{x_{ij}\}) \propto p(\sigma) \prod_{i=1}^{p} \frac{1}{t_i} f_{n_i - 1} \left[\frac{\bar{x}_{i.} - \mu}{t_i}; \frac{2\sigma^2}{t_i^2} \right].$$

- The posterior of μ given $\sigma = 0$ is a product of scaled t-densities centered at the x_i, since

$$\frac{1}{t_i} f_{n_i - 1} \left[\frac{\bar{x}_{i.} - \mu}{t_i}; 0 \right] = \frac{1}{t_i} T'_{n_i - 1} \left(\frac{\bar{x}_{i.} - \mu}{t_i} \right).$$

- We will take $p(\sigma) = 1$, though an arbitrary proper prior does not introduce additional difficulties.

Marginals for μ and σ:
Arsenic Data

Marginal Posterior: Apple Fiber Data

Small Simulation Comparing
Bayesian and Frequentist Intervals

$$
\begin{aligned}
\mu &= 0 \\
\sigma_i &= \sigma_e \\
\sigma^2 + \sigma_e^2 &= 1 \\
\rho &= \sigma^2/(\sigma_e^2 + \sigma^2) = 1/2
\end{aligned}
$$

A Two-Way Mixed Model
(Heteroscedastic, no Interaction)

$$x_{ijk} = \theta_k + \delta_i + e_{ijk},$$

- $i = 1, \ldots, p$ Laboratories

- $j = 1, \ldots, n_i$ Replicates

- $k = 1, \ldots, m$ Materials

$$\delta_i \sim \mathsf{N}\left(0, \sigma^2\right)$$

$$e_{ijk} \sim \mathsf{N}\left(0, \sigma_i^2\right)$$

Some notation: $\tau_i^2 \equiv \sigma_i^2/(n_i m)$, $\nu_i \equiv n_i m - 1$.

ML Equations

$$\theta_k - \bar{\theta} \equiv \phi_k = \frac{\sum_{i=1}^{p}(\bar{x}_{i \cdot k} - \bar{x}_{i \cdot \cdot})/\tau_i^2}{\sum_{i=1}^{p} 1/\tau_i^2}$$

$$\bar{\theta} = \frac{\sum_{i=1}^{p} \gamma_i \bar{x}_{i \cdot \cdot}}{\sum_{i=1}^{p} \gamma_i}$$

$$\sigma^2 = \frac{\sum_{i=1}^{p} \gamma_i \left[(\bar{x}_{i \cdot \cdot} - \bar{\theta})^2 + \frac{\nu_i t_i^2}{1 - \gamma_i} \right]}{\sum_{i=1}^{p} n_i}$$

Where $\tau_i^2 \equiv \sigma_i^2/(n_i m)$, $\nu_i \equiv m n_i - 1$, $\gamma_i \equiv \sigma^2/(\sigma^2 + \tau_i^2)$, and

$$t_i^2 \equiv \frac{\sum_{j,k}(x_{ijk} - \bar{x}_{i \cdot k})^2 + n_i \sum_k (\bar{x}_{i \cdot k} - \bar{x}_{i \cdot \cdot} - \phi_k)^2}{\nu_i n_i m}$$

ML Equations (Cont'd)

The weights $\{\gamma_i\}_{i=1}^p$ are roots of the cubic equations

$$
\begin{aligned}
&\gamma_i^3 - (a_i + 2)\gamma_i^2 + \\
&\left[(n_i m + 1)a_i + \nu_i b_i + 1\right]\gamma_i - \\
&n_i a_i = 0
\end{aligned}
$$

where

$$
a_i \equiv \frac{\sigma^2}{(\bar{x}_{i..} - \bar{\theta})^2}
$$

and

$$
b_i \equiv \frac{t_i^2}{(\bar{x}_{i..} - \bar{\theta})^2}.
$$

An ML Iteration

1. Begin with estimates $\left\{\gamma_i^{(s)}\right\}$.

2. Calculate the following:

$$\phi_k^{(s+1)} = \frac{\sum_{i=1}^p (\bar{x}_{i\cdot k} - \bar{x}_{i\cdot\cdot})/\tau_i^{2(s)}}{\sum_{i=1}^p 1/\tau_i^{2(s)}}$$

$$\bar{\theta}^{(s+1)} = \frac{\sum_{i=1}^p \gamma_i^{(s)} \bar{x}_{i\cdot\cdot}}{\sum_{i=1}^p \gamma_i^{(s)}}$$

$$\sigma_{(s+1)}^2 = \frac{\sum_{i=1}^p \gamma_i^{(s)} \left[(\bar{x}_{i\cdot\cdot} - \bar{\theta})^2 + \frac{\nu_i t_i^2}{1 - \gamma_i^{(s)}} \right]}{\sum_{i=1}^p n_i}$$

3. Note that if the ϕ_k are constrained to satisfy the above ML equation, then

$$t_i^2 = \frac{\sum_{j,k} (x_{ijk} - \bar{x}_{i\cdot\cdot})^2 - \sum_k \phi_k^2/m}{n_i \nu_i m}$$

4. Solve the cubics for new estimates $\gamma_i^{(s+1)}$, and iterate.

Some Theoretical Results for Two-Way Mixed Model

The one-way results discussed earlier generalize:

- Monotone convergence

- All stationary values of likelihood in box in $(\mu, \sigma, \sum_k \phi_k^2)$ space.

- Exactly one weight $\gamma_i \in [0, 1]$, unless ith lab an outlier and n_i small

- Variances cannot be negative at solution to likelihood equation.

Hierarchical Model With Noninformative Priors: Two-Way Model

$i = 1, \ldots, p$ indexes laboratories

$j = 1, \ldots, n_i$ indexes measurements

$k = 1, \ldots, m$ indexes materials

$$
\begin{aligned}
p(x_{ijk}|\delta_i, \theta_k, \sigma_i^2) &= \mathsf{N}\,(\delta_i + \theta_k, \sigma_i^2) \\
p(\sigma_i) &\propto 1/\sigma_i \\
p(\delta_i|\mu, \sigma^2) &= \mathsf{N}\,(\mu, \sigma^2) \\
p(\theta_k) &\propto 1 \\
p(\sigma) &\propto 1
\end{aligned}
$$

Posterior of (μ, σ):
Two-Way Model

- The posterior of $(\{\theta_k\}, \sigma)$ is

$$p(\{\theta_k\}, \sigma | \{x_{ijk}\}) \propto p(\sigma) \prod_{i=1}^{p} \frac{1}{t_i} f_{\nu_i} \left[\frac{\bar{x}_{i..} - \bar{\theta}}{t_i}; \frac{2\sigma^2}{t_i^2} \right].$$

where $f_\nu(\cdot, \theta)$ is the *generalized t-distribution* defined earlier, and

$$t_i^2 \equiv \frac{\sum_{j,k}(x_{ijk} - \bar{x}_{i\cdot k})^2 + n_i \sum_k (\bar{x}_{i\cdot k} - \bar{x}_{i..} - \phi_k)^2}{\nu_i n_i m}$$

Summary

- Interlaboratory studies are important in many fields.

- W.J. Youden and John Mandel of the National Bureau of Standards have left an important mark on the simple methods in common use today, through the Youden plot, the Row-Linear Model, and other ideas.

- But there remains considerable opportunity for new methodology, more realistic and computationally intensive.

- The new results presented here include
  - Relating an ad-hoc procedure due to Mandel and Paule to REML
  - New results for the one-way heteroscedastic model, including a way to find all statationary points of the likelihood, and to easily calculate profile likelihoods.
  - A Bayesian hierarchical model leads to approximate confidence regions, and promises to be useful, with an informative prior on σ, when the number of labs. is small.